Short Research Article

Is the Emotional Dog Blind to Its Choices?

An Attempt to Reconcile the Social Intuitionist Model and the Choice Blindness Effect

Marek A. Vranka[1] and Štěpán Bahník[2]

[1]Department of Psychology, Faculty of Arts, Charles University, Prague, Czech Republic
[2]Department of Psychology, University of Würzburg, Germany

Abstract: Previous choice blindness studies showed that people sometimes fail to notice when their choice is changed. Subsequently, they are willing to provide reasons for the manipulated choice which is the opposite of the one they made just seconds ago. In the present study, participants first made binary judgments about the wrongness of described behaviors and then were shown an opposite answer during a second reading of some of the descriptions. Half of the participants saw the answer during the second presentation of the description and the other half saw it only after the presentation. Based on Haidt's Social intuitionist model, we hypothesized that participants in the latter group would be less likely to reconcile their intuition with the presented answer and thus they would be more likely to reject it. However, we found no difference between the groups.

Keywords: choice blindness, social intuitionist model, moral judgment, rationalization, intuition, dual-process models

People are willing to spend hours arguing about moral issues and persuading each other about the rightness of their political opinions. Yet, people were easily manipulated into defending an opinion opposite to the one they had stated just a moment ago when Hall, Johansson, and Strandberg (2012) reversed their answers in an opinion poll using a simple magic trick. The study employed the choice blindness paradigm (Johansson, Hall, & Sikström, 2008) which shows that people can often fail to notice a change in the choice they have made just seconds before. In the study of Hall et al. (2012), for example, after agreeing that purchasing sexual services is morally defensible, some participants failed to notice that the question wording was reversed during the second presentation, so their original answer now expressed the opposite opinion. Moreover, many people had no trouble arguing for the reversed position, suggesting that the effect cannot be explained by a simple lack of focus during the second presentation (Hall et al., 2012).

Similar choice blindness effects have been observed in various other settings, including consumer choice of jams and teas (Hall, Johansson, Tärning, Sikström, & Deutgen, 2010), judgment of attractiveness (Johansson, Hall, Sikström, & Olsson, 2005), rating of pleasantness of voices (Sauerland, Sagana, & Otgaar, 2013), and evaluation of political issues (Hall et al., 2013).

The phenomenon of choice blindness in the moral domain, and especially the participants' willingness to defend a reversed answer, seems to support the current theories of moral judgment that emphasize people's ability to post hoc rationalize judgments (Haidt, 2001).[1] The most prominent of this type of theories, Haidt's Social intuitionist model, proposes that when people judge whether something is morally right or wrong, they usually base their judgment on fast, automatic intuitions associated with positive or negative feelings (Haidt, 2007). Rational moral reasoning, traditionally considered as the primary source of moral judgments (Haidt, 2008; Kohlberg, 1976), comes to play only after the initial judgment is formed. Its function is mainly social and lies in communicating and defending one's moral judgment to others (Haidt, 2001). Moral reasoning is therefore selectively biased in favor of finding evidence and reasons supportive of the initial intuitive judgments (Ditto, Pizarro, & Tannenbaum, 2009). This limited function of moral reasoning can be most clearly illustrated in the phenomenon of moral dumbfounding – a situation when people express strong moral judgment even

[1] In Hall et al. (2012), 53% of participants at least once argued unequivocally for a moral attitude opposite of the one they had held originally.

DOI: 10.1027/1618-3169/a000325

when they are unable to provide sufficient reasons for its support. A common demonstration of moral dumbfounding uses Haidt's Mark and Julie incest scenario, in which siblings have a safe sexual intercourse without any negative emotional consequences (Haidt, Bjorklund, & Murphy, 2000). Even after participants admit that they are unable to say why the act was wrong, most of them continue to judge it as immoral, presumably because of their initial negatively laden intuition (but see Royzman, Kim, & Leeman, 2015, for a critique of this interpretation).

However, as Hall et al. (2012) note, the choice blindness phenomenon also poses a problem for the Social intuitionist model – when people are presented with the same moral issue for the second time only a few moments later, they should have the same intuition as after the first reading. This intuition should then be in conflict with the reversed answer shown to them. How is it then possible that people accept the reversed judgment?

To make the situation more concrete, imagine Mary who reads a short scenario about a woman that cheated on her husband, who was too unaffectionate and not sensitive to her needs. Mary judges the behavior as wrong. After a while, she reads the same scenario again; however, now she is told that her previous answer was "not wrong." Her intuitive reaction should presumably be the same as the first time – she should find the described behavior wrong. Nevertheless, influenced by the manipulated answer, she now judges it as not wrong.

One can hypothesize that after seeing the presented answer, she started searching for reasons supporting it, in order to stay consistent and prevent cognitive dissonance (Albarracin & Wyer, 2000). Or maybe she simply accepted the presented judgment on the assumption that she must have had good reasons to make such a judgment after the first reading. Both explanations assume that the initial intuition can be easily overridden by reasoning – either by finding reasons for the presented answer or by believing that such reasons exist. However, Haidt's model states that a subsequent modification of intuition by deliberative reasoning is supposed to be difficult and rare (Haidt, 2001). The existence of the choice blindness effect then suggests two possible consequences for the Social intuitionist model: either the Social intuitionist model can be somehow reconciled with the choice blindness effect, or the model should be abandoned in favor of one that is more readily able to accommodate the effect.

Returning to our example, a possible explanation consistent with the Social intuitionist model is that when judging the woman for the second time, Mary not only intuitively judged her behavior, but also automatically started searching for support of the manipulated answer as soon as she saw it. In result, Mary's judgment was shifted in the direction of the presented answer. Whether the shift away from the original intuitive judgment was sufficient depended on the strength of the original intuition and ease of the search for support of the presented answer. In our example, Mary might have focused on the emotional distance of the husband and felt that the marriage was not strong anyway. This provided her with enough intuitive support for not condemning the woman's behavior. In consequence, her new judgment was in accord with the manipulated answer.

The Social intuitionist model describes a mechanism how a judgment of one person can influence intuition and subsequent judgment of someone else (Haidt, 2001) – the so-called social persuasion link. The model can be therefore reconciled with the choice blindness effect if we assume that seeing the manipulated answer influences the intuition that is being formed similarly as being aware of someone other's judgment. If that is the case, we would expect that the probability of rejecting the manipulated answer should be higher when it is presented later, after the judged scenario has been completely read, than when it is presented together with it.[2] On the other hand, if presenting the manipulated answer only after the scenario has been read does not affect the occurrence of choice blindness, it would suggest that initial intuitions can be subsequently overridden more easily than Haidt argues.

In the present experiment, we manipulated the time of presentation of the previous choices to test the described possibility of reconciliation of the choice blindness effect and the Social intuitionist model. While previous studies of choice blindness in the moral domain were interested in abstract moral principles and used response scales, the present study used a binary choice and the moral judgment was about descriptions of real-world moral transgressions.

Methods[3]

Participants

We recruited 201 participants (72% female) from our laboratory subject pool.[4] The experiment was conducted

[2] Note that the argument is probabilistic and it should therefore apply even if some people decide rationally (see Haidt, 2003) or if the shift away from the original intuitive judgment is sometimes not sufficient to make people accept the manipulated answer.

[3] Materials, data, and R scripts used for analysis can be found on https://osf.io/q72a5/.

[4] The sample size was preregistered (https://osf.io/veus3/). Given the complexities of estimating statistical power of multilevel models, we chose a conservative approach and aimed for 200 participants, that is, a sample size usually large enough to have reasonable power for small-to-middle sized effects.

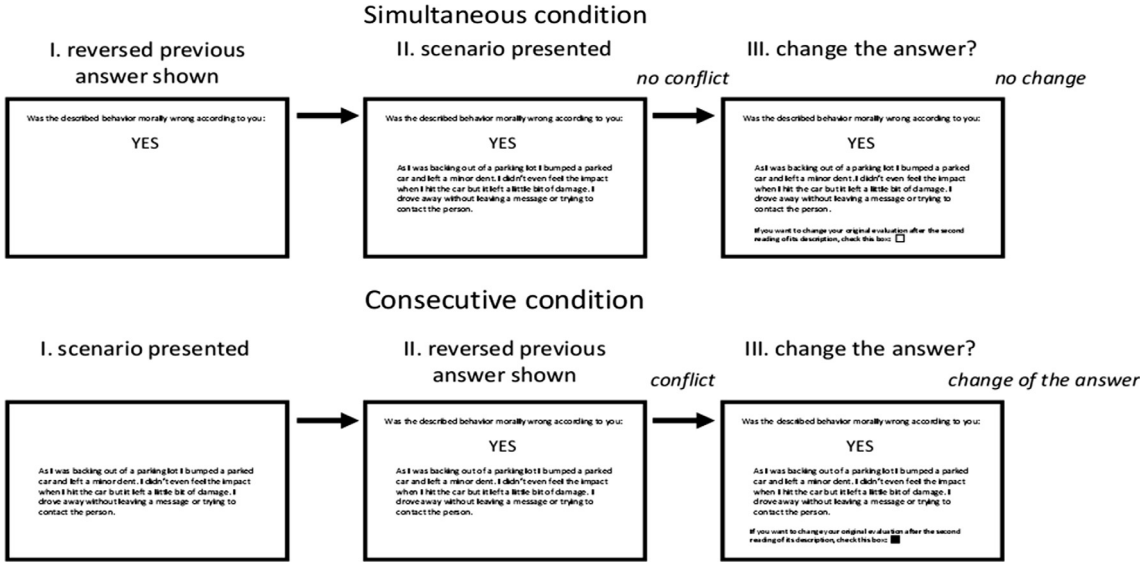

Figure 1. Schema of the second presentation of scenarios – an example of a trial with a manipulated initial answer. One group of participants was shown manipulated answers before the scenario was presented (see Figure 1, top). The other group was presented the scenario first (Figure 1, bottom). In this illustration we assume that the initial answer was NO. Once both the answer and whole scenario were presented, participants were given a chance to change their initial answer. They were also asked whether they initially decided rationally or intuitively (not shown). According to the Social intuitionist model, we hypothesized that reading the scenario for the second time should lead to the same intuition as it did the first time, which would then conflict with the presented manipulated answer. This would likely lead to a change of the answer (bottom, in italics). On the other hand, participants should be more likely to retain the manipulated answer if it is shown before the presentation of the scenario. Since the intuition would be modified in accord with the manipulated answer and thus consistent with it, there would be no conflict of the manipulated answer with the intuition (top, in italics).

in sessions of up to 14 participants. The median age of the participants was 22 years and 80% of them were university students.

Procedure

The experiment was administered in a laboratory setting, on computers using a custom written Python program. The experiment was conducted in Czech.

Participants were given descriptions of 40 real-world situations where a person behaves in a way that can be considered morally wrong (e.g., "As I was backing out of a parking lot I bumped a parked car and left a minor dent. I didn't even feel the impact when I hit the car but it left a little bit of damage. I drove away without leaving a message or trying to contact the person." or "My first husband was never sensitive to my needs. I cheated on him because he wasn't there for me. He was not affectionate. I decided to have an affair because I felt like he deserved it."). The situations were randomly chosen for each participant from a total of 67 situations which we selected from a larger set by Knutson et al. (2010). The situations were displayed one-by-one in a random order. Each situation was presented at a speed of 65 ms per letter and additional 350 ms after each line (i.e., slightly slower than the normal reading speed). Once the whole text was displayed this way, participants answered a yes-no question asking whether they think that the described behavior was morally wrong and then moved to the next situation.

After judging all 40 situations, participants were once again presented with all scenarios (in the same order and in the same way as during the first presentation). Importantly, participants were randomly divided in two groups. One group was shown the answers from the first presentation right from the moment the description started to be displayed (*simultaneous condition*). The other group was shown the answers only after the description was displayed in full (*consecutive condition*). Five randomly selected initial answers out of the 40 were reversed during the second presentation (from "Yes" to "No" and vice versa). See Figure 1 for a schema illustrating the manipulation. During this presentation, participants were asked whether they initially decided intuitively or rationally, which provided an overt reason for the repeated presentation of the situations. Participants were also given an opportunity to indicate that they want to change their initial answer after the second reading ("If you want to change your original evaluation of the behavior after the second reading of its description, check this box"). The decision to change the answer served as the primary dependent variable.

After the experiment, we asked participants whether they noticed anything unexpected during the experiment and if

so, what did they notice. Then we explained the manipulation we used and asked them whether they had noticed it. We used these measures of noticing the manipulation only for exploratory analyses – although we expected participants in the consecutive condition to notice the manipulation more often, we considered this question only as a crude measure for examining our main hypothesis. The decision to change the presented answer described above was available for every trial and thus it allowed for an analysis with much higher statistical power. Additionally, we asked several questions that we used for exclusions of participants who were not focused or knew the goal of the study. Full wording of all questions can be found on https://osf.io/7p3u2/wiki/.

Results

Following preregistered exclusion criteria, we excluded 10 participants who indicated that we should not use their data (because they did not understand instructions or were not focused during the session) and 1 participant who both noticed the manipulation of the answers and knew what choice blindness means. All analyses were done with the data of the remaining 190 participants. Analysis was conducted using mixed-effect models (Baayen, Davidson, & Bates, 2008; Gelman & Hill, 2006).

Prevalence of Choice Blindness

Participants indicated that they wanted to change their initial answer in 54% of the manipulated trials (516 out of 950) in comparison to only 3% of the non-manipulated trials (199 out of 6,650). The low percentage of changed non-manipulated trials shows that people indicated that they want to change the answer almost only if they noticed the manipulation. Notably, 85% of participants changed at least one manipulated answer but only 14% of participants changed all five manipulated answers. In the debriefing, 45% of participants reported that they noticed something unexpected during the experiment. Out of these participants, 88% provided an open answer suggesting that they noticed the manipulation. Finally, 81% of all participants reported that they noticed the manipulation when it was described to them. In general, the rate of choice blindness (i.e., the proportion of manipulated trials not detected) seemed to be comparable with previous studies (Hall et al., 2012; Johansson et al., 2008; Sagana, Sauerland, & Merckelbach, 2014).

Confirmatory Analysis

We used mixed-effect logistic regression to test the primary hypothesis. The change of the presented answer on the second reading served as the dependent variable.[5] Situations and participants were treated as random factors and the effects of experimental condition and manipulation of an answer were included as predictors. We used both varying intercept and varying slope for the effect of the manipulation for situations and only varying intercept for participants.

As described above, participants were more likely to change the presented answer when it was manipulated than when it was not manipulated, $z = 30.60$, $p < .001$, OR = 57.61, 95% CI = [44.43, 74.69].[6] Participants in both conditions were equally likely to change their answers, $z = -0.58$, $p = .57$, OR = 0.91, 95% CI = [0.66, 1.25]. Most importantly, there was no significant interaction between condition and manipulation of an answer, $z = -0.06$, $p = .96$, *ratio of OR* = 0.99, 95% CI = [0.65, 1.50].

We were mainly interested in the effect of condition on the manipulated trials. Therefore, we conducted a mixed-effect logistic regression only with the manipulated trials. We did not find any difference in the rate of change of an answer between the two conditions. The overall rate of changed answers was 53.2% for the simultaneous condition and 55.2% for the consecutive condition, $z = -0.46$, $p = .65$, OR = 0.89, 95% CI = [0.54, 1.47]. The results did not support our hypothesis since there was no difference in the occurrence of choice blindness based on the time of presentation of the initial answer.

Exploratory Analysis

While we found no effect of participant's condition on the probability of changing an answer in the manipulated trials, it is possible that there were other factors that influenced it. Therefore, we conducted analyses exploring effects of various predictors of the change of an answer. These analyses are not related to the main aim of the study, but they

[5] It is also possible to analyze change of an answer compared to the original answer, that is, whether participants retained the presented answer on non-manipulated trials and changed it on manipulated ones. However, we were interested mainly in the probability of changing the presented answer on manipulated trials and not in the stability of judgment and our chosen measure is more appropriate for this.

[6] Note that the odds ratio does not correspond to the percentages exactly, because it takes into account variance explained by other factors.

can shed more light on the phenomenon of choice blindness. We did not include the experimental condition variable in them since it did not affect the rate of changing the answers and it was not of interest for the analyses.

First, we analyzed the effect of the number of previous manipulated trials. The interaction of the manipulation and number of previous manipulated trials was significant, $z = 4.60$, $p < .001$, *ratio of OR* = 3.47, 95% CI = [2.04, 5.90]. Follow-up separate analyses for manipulated and non-manipulated trials showed that in later trials participants were more likely to change a manipulated answer, $z = 3.64$, $p < .001$, OR = 2.22, 95% CI = [1.45, 3.41], but less likely to change a non-manipulated answer, $z = -2.89$, $p = .004$, OR = 0.59, 95% CI = [0.42, 0.85].

It is also possible that wrongness or unambiguity of the judgment of the described behavior might have influenced the probability of change of an answer. For each situation we computed a wrongness rating as the proportion of people who considered the described behavior wrong. Then we computed a measure of unambiguity of wrongness by taking the distance between the proportion of 0.5 and the wrongness rating. The interaction of the manipulation and unambiguity of wrongness was significant, $z = 5.10$, $p < .001$, *ratio of OR* = 9.15, 95% CI = [3.91, 21.41]. Including the wrongness measure did not improve the model, $X^2(2) = 0.47$, $p = .79$. Separate analyses for manipulated and non-manipulated trials showed that participants were more likely to change their answers in more unambiguous manipulated trials, $z = 3.68$, $p < .001$, OR = 3.46, 95% CI = [1.79, 6.69], but less likely to change their answers in unambiguous non-manipulated trials, $z = -3.23$, $p = .001$, OR = 0.34, 95% CI = [0.18, 0.65]. This shows that manipulations in situations that were clearly wrong or permissible were more likely to be changed to the original answer.

The effect of unambiguity of wrongness might be related to the time spent on the initial answer as situations that were clearly wrong or clearly permissible were judged faster, $t(65.2) = -4.31$, $p < .001$, $b = -4.06$, 95% CI = [−5.90, −2.22].[7] From this perspective, longer duration spent on the initial answer would be negatively associated with the probability of changing a manipulated answer. On the other hand, it can be argued that deeper processing of materials should be associated with a higher probability of noticing the manipulation. The results were more consistent with the former possibility; a longer time to come up with the initial answer was negatively associated with the probability of changing a manipulated answer, $z = -1.85$, $p = .06$, OR = 0.98, 95% CI = [0.96, 1.00], and the effect was in the opposite direction for non-manipulated trials. That is, participants were more likely to change their non-manipulated previous answer when it took more time, $z = 3.98$, $p < .001$, OR = 1.02, 95% CI = [1.01, 1.03].

However, it is possible that the time spent on the initial answer may be also related to the number of previous manipulated trials. That is, the effect of the time spent on the initial answer may be driven by the decreasing time spent answering the initial question with subsequent trials; $t(7,356.1) = -6.80$, $p < .001$, $b = -0.05$, 95% CI = [−0.07, −0.04], for the effect of the order of a trial on the time spent answering the initial question (in seconds). While adding the time spent on the initial answer to the model with the number of previous manipulated trials does not result in a significantly better fit for manipulated trials, $X^2(1) = 2.15$, $p = .14$, it does for non-manipulated trials, $X^2(1) = 14.29$, $p < .001$, as well as for the model including both types of trials, $X^2(2) = 14.99$, $p < .001$. This suggests that the effect of the time spent on the initial answer might have influenced the rate of changes of manipulated answers only because it correlated with the trial order. However, for non-manipulated trials, longer time spent on the initial answer probably indicated more uncertainty about the decision rather than deeper processing and it was thus associated with a higher probability of change after the second reading, $z = 4.04$, $p < .001$, OR = 1.02, 95% CI = [1.01, 1.03].

Finally, we analyzed whether participants were more likely to change the presented answer if they initially answered that the described behavior was wrong. Since 58% of the initial answers were that the behavior was wrong, we also included the unambiguity of wrongness as a covariate. The interaction of the previous answer and manipulation was significant, $z = 7.87$, $p < .001$, *ratio of OR* = 5.83, 95% CI = [3.76, 9.05], so we conducted separate analyses for manipulated and non-manipulated trials. The analysis for manipulated trials showed that participants were more likely to change the manipulated answer if they initially considered the behavior wrong, $z = 4.25$, $p < .001$, OR = 2.14, 95% CI = [1.51, 3.05]. The effect of wrongness unambiguity was reduced, but remained significant, $z = 2.63$, $p = .009$, OR = 2.50, 95% CI = [1.26, 4.95]. On the other hand, participants were less likely to change their non-manipulated answers if they previously answered that the behavior was wrong, $z = -7.00$, $p < .001$,

[7] The time spent on the initial answer was measured in seconds and the median time of the answer was 4.1 s ($M = 7.0$ s, $SD = 8.7$ s). Note that the short time spent on the initial answer is caused by the method of presentation of situations during which people had time to form their judgment. The short time also shows that the duration of the presentation of a situation is usually sufficient to form a judgment and participants require more time to make the decision only in a minority of cases. This suggests that during the second reading of scenarios, participants in the consecutive condition usually have their intuitions already formed before seeing the presented answer.

OR = 0.30, 95% CI = [0.22, 0.42], and the effect of unambiguity of wrongness was not significant, $z = -1.48$, $p = .14$, OR = 0.57, 95% CI = [0.28, 1.20].

Discussion

On the one hand, the existence of choice blindness in the moral domain seems to support Haidt's (2001) Social intuitionist model by demonstrating that arguments in favor of a moral judgment are often constructed through post hoc confabulation (Hall et al., 2012). On the other hand, choice blindness also poses a problem for Haidt's model – it is unclear how can people accept a reversed answer if they have the same intuition on both presentations of a given moral issue and this intuition is resistant to subsequent modifications, as Haidt's model assumes. Our goal was to examine one possible explanation that could reconcile the choice blindness effect and the Social intuitionist model. Namely, we hypothesized that seeing a manipulated answer during the second presentation modifies the intuition, which becomes more in accord with the answer, as suggested by the social persuasion link in Haidt's model. Therefore, we expected that presenting the manipulated answer only after the judged scenario has been read would increase the probability of rejecting it. However, the results showed no support for our hypothesis. We found no effect of the time of presentation of the manipulated answer on the probability of changing it back to the initial one.

Although there are several alternative explanations that could account for the lack of the observed difference without rejecting Haidt's model of moral judgment, we believe our results strongly suggest that subsequent modifications of the initial intuitions occur more frequently than the Situational intuitionist model proposes.

As in most choice blindness studies, we have assessed only whether participants rejected the presented answer or not. Therefore, it is difficult to discern when participants really accepted a reversed choice as their own and when they just failed to communicate that they did not consider it as their own. One could then object that the two conditions in our study differed in the rate of noticing the manipulation, but participants were unwilling to reject the manipulated answer (e.g., because they were afraid it could ruin the experiment), which masked the difference between the conditions. Hall et al. (2012) overcame this possible objection by asking participants to argue for the presented stance on a given moral issue. It was thus possible to assess whether the provided arguments aligned with the manipulated or original answer. Although we did not ask for arguments in favor of the presented answer, we explicitly asked participants whether they wanted to change it. We believe that giving participants this option was sufficient to eliminate their reservations about rejecting the presented answer when they thought that it contradicted their judgment. The fact that 85% of participants rejected at least one manipulated answer further supports the assumption that participants had no reservations about rejecting the presented answer when they felt that it was not in accord with their judgment.

Another possible objection is that participants did not pay sufficient attention to the presented answers. If that was the case, observed acceptance of manipulated answers would be meaningless. Moreover, if participants ignored the presented answers, it would not matter when they were presented. This would mean that both experimental conditions were virtually identical and no difference between them could have been expected. However, the possibility that participants ignored the presented answers seems unlikely. We displayed each answer in bold capital letters in a central location of the screen. We also brought participants' attention to the presented answer by asking them whether they want to change it. We also excluded participants who stated in the debriefing that they were not focused during the study. Furthermore, only 15% of participants retained all five manipulated answers, suggesting that the majority of participants examined the presented answer closely at least on some trials. In addition, previous research (Hall et al., 2012; Johansson et al., 2008) shows that participants usually do not have difficulty to provide an explanation of a manipulated choice for which they clearly need to be aware of it. Therefore, the objection based on insufficient attention does not seem to be warranted.

Alternatively, it is possible that participants did not have a strong intuition about wrongness of described behaviors and decided mostly at random when they were first presented with the descriptions. Because both options seemed equally acceptable, they did not indicate that they want to change the displayed answer. This possibility is in accord with the finding that the unambiguity of wrongness was related to the probability of rejecting the presented answer as well as with the finding by Hall et al. (2012) that manipulations of trials with more extreme judgments are more often noticed and corrected. However, both in our study and in the study by Hall et al., participants sometimes failed to notice manipulations even in issues that were the least ambiguous. The much lower rate of changes in the nonmanipulated trials compared to manipulated trials in the present study further speaks against this possibility. Moreover, Sagana et al. (2014) found that even participants who failed to notice a manipulation could reproduce their initial answer correctly when asked later. Therefore, it seems unlikely that a lack of a clear moral view can account for the absence of difference between the experimental conditions in the present study.

Another possible objection could be based on the fact that presenting a manipulated answer immediately after displaying the whole scenario does not provide enough time for an intuition to form. If that was the case, our two experimental conditions would not really differ. However, the scenarios were displayed letter-by-letter at a speed slightly slower than the average reading speed, therefore giving participants enough time to not only read, but also to simultaneously process the described situations. This is supported by the relatively fast median response time of 4 s (which includes selection of a button corresponding to an answer) observed during the first presentation of the scenarios.

The present study was not able to reconcile the choice blindness effect with the Social intuitionist model as it aimed to do. The results are more in line with theories of moral judgment which assume that moral intuitions can be overridden by subsequent reasoning and deliberation (Cushman, Young, & Hauser, 2006; Paxton & Greene, 2010; Pizarro & Bloom, 2003). For these theories, it is not difficult to account for the choice blindness effect as well as for the lack of observed difference between our experimental conditions. For example, Greene's dual-process model of moral judgment assumes that an initial intuition can be overridden by deliberative considerations (Greene et al., 2009).[8] According to Greene's model, people can read a scenario, form an intuition about it, and then suppress the intuition when they are later shown the manipulated answer. It would then be possible that the manipulated answer can be retained with the same probability regardless of whether it was presented before or after the intuition formation.

However, even from the viewpoint of dual-process models, it could be argued that participants should be more likely to retain the manipulated answer when it is presented from the beginning. The argument would be that the intuition is modified in accord with the presented answer and its subsequent overriding might be easier or even unnecessary. We agree that this is at least theoretically possible. On the other hand, dual-process models do not predict precisely how difficult the overriding of intuitions actually is. In case it is relatively easy, it might have only limited impact whether the manipulated answer is presented together with or after the scenario.

Nevertheless, it is still possible that if the manipulated answer was presented with a longer delay after displaying the scenario, the rate of choice blindness might have been lower. We can assume that without being shown the manipulated answer one would have the same intuition as well as the same or similar deliberations during the second presentation of the scenario. This would lead to the same judgment as was made during the first presentation. If this finalized judgment is only then confronted with the manipulated answer, it should be more difficult to accept the manipulated answer than it would have been if it had been presented during earlier stages of the judgment process. A future study could compare the choice blindness rate in the simultaneous condition and a consecutive condition with a longer delay between presentation of the scenario and manipulated answer. This might help to identify crucial factors for occurrence of choice blindness. Namely, whether the information about the previous choice has to be presented before the second judgment is definitively made.

Our results support findings of previous studies demonstrating that moral intuitions are more easily and frequently modified than Haidt's model assumes. For example, Pizarro, Uhlmann, and Bloom (2003) showed that participants assigned different levels of blame to perpetrators of moral transgressions when they were asked to make rational, objective judgments instead of intuitive, gut judgments. This result can be interpreted as evidence either of an effect of one's reasoning on moral judgment or of an effect of the experimenters' instructions on the judgment of the participants. Haidt's model would have difficulties to accommodate either of the effects. However, this particular design is prone to a potentially strong effect of experimenter demand which makes it unclear whether the judgment was affected or whether participants just modified their answers in line with experimenter demand and contrary to their judgment (Paxton & Greene, 2010). Our study is less susceptible to the experimenter demand effect, as discussed above. Therefore, the present results further corroborate the results of Pizarro et al. (2003).

In their review, Paxton and Greene (2010) mention various studies that suggest a more prominent role of moral reasoning in moral judgment. However, virtually all of them are vulnerable to the objection that the studied factors could directly affect moral intuitions, making all findings compatible with the intuitive account of moral judgment and Haidt's model. For example, participants in Wheatley and Haidt's (2005) study were hypnotized to feel a flash of disgust when reading a specific word. Afterwards, when participants read the disgust inducing word in a scenario describing a completely innocent behavior, they judged it as more immoral than participants who read a version

[8] Although intuition is supposed to be based on deontological rules and deliberation is supposed to be motivated by utilitarian concerns according to Greene, this view is criticized as overly simplistic (Kahane, 2012). Furthermore, a recent study (Royzman, Landy, & Leeman, 2015) shows that higher reflective ability is not associated with strictly utilitarian judgments and lower ability with deontological judgments. The findings cast further doubts on Greene's intuitive-deontological and deliberative-utilitarian pairings.

without the word. Paxton and Greene (2010) note that the majority of participants who felt a flash of disgust during the reading still judged the behavior as innocent – therefore they probably had to take into account that there was nothing objectively wrong with the behavior when making the judgment. But they were aware of the fact that there is nothing objectively wrong with the behavior from the moment they started judging it, so their judgment could have been fully intuitive. In a similar vein, when people judge moral dilemmas under cognitive load or time pressure, they condemn utilitarian decisions more often than they do under control conditions (Trémolière, De Neys, & Bonnefon, in press). Although these results are interpreted as evidence of the role of reflective processes in moral judgment, it is again possible that people who are under cognitive load or time pressure just have different intuitions. For example, they could experience more intensive distress that they associate with stronger condemnation of the utilitarian behavior. On the other hand, in the consecutive condition in our study, the information about the reversed previous answer was presented only after the whole scenario was read and the intuition was formed. Presumably, this intuition was identical with the one formed during the first presentation of the scenario. Despite the presence of the same intuition, participants still retained 46% of the reversed answers, so the willingness to retain the reversed judgment has to be explained by subsequent deliberation.

Beside the test of the main hypothesis, our study also contributes to the broader choice blindness literature. We supported the previous finding that the unambiguity of a judgment is positively related to the probability of rejecting a manipulated choice (Hall et al., 2012). Participants were also more likely to reject a manipulated answer if they initially answered that the described behavior was wrong. It is possible that it was easier to find support for condemnation of the described behavior when it was initially considered right than to justify it when it was considered wrong. We also found that after the rejection of one manipulation, subsequent manipulations were rejected with a higher probability. Furthermore, we demonstrated the choice blindness effect at a rate comparable to previous studies, even though we used binary moral judgment instead of judgment on a scale and computer administration instead of administration by a researcher, therefore eliminating a possible experimenter effect. Moreover, we offered participants an option to directly indicate that they want to change the presented answer. In the previous studies, participants usually had to actively announce that they disagree with the presented choice, which could have prevented some of them from revealing that they did not truly accept the manipulated answer. All of these differences could be expected to reduce the number of retained manipulated answers.

On the other hand, we used judgments about common moral transgressions instead of judgments about general moral principles and issues with a great social impact. It is therefore possible that participants in our experiment held less extreme opinions about the objects of their judgments. Furthermore, participants judged 40 situations in our study and they saw manipulated answers only after judgment of all the situations. The task could have been therefore more cognitively taxing. In combination with the longer delay between the two presentations, it might have increased the rate of choice blindness, possibly working against the previously listed factors.

In summary, there are two main contributions of our study: We have built on Hall et al. (2012) and examined one possible way how to reconcile the choice blindness effect with the Social intuitionist model of moral judgment, but we failed to find supporting evidence for it. Our results therefore strengthen the original objection by Hall et al. (2012), further questioning the proposed prominent role of intuition in moral judgment. We have also used the choice blindness paradigm as a tool that enabled us to infer otherwise hardly accessible intuitive mental processes. As far as we know, our attempt constitutes one of the first such applications of the choice blindness effect.

Acknowledgments

We would like to thank Pan Houdek for his help with the recruitment of participants, Anand Krishna for a language check, and our research assistants Laura Bechyňová, Nikola Frollová, Petra Karbanová, and Šimon Kucharský for their help with preparation of the study materials. The work of Marek Vranka was funded by Charles University in Prague (Specific university research at the Faculty of Arts VG2015 and PRVOUK Grant No. 03). The work of Štěpán Bahník was funded by Deutsche Forschungsgemeinschaft (DFG-RTG 1253/2). The authors contributed equally to this work.

References

Albarracin, D., & Wyer, R. S. Jr. (2000). The cognitive impact of past behavior: influences on beliefs, attitudes, and future behavioral decisions. *Journal of Personality and Social Psychology, 79*, 5–22.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*, 390–412.

Cushman, F., Young, L., & Hauser, M. (2006). The role of conscious reasoning and intuition in moral judgment testing three principles of harm. *Psychological Science, 17*, 1082–1089.

Ditto, P. H., Pizarro, D. A., & Tannenbaum, D. (2009). Motivated moral reasoning. *Psychology of Learning and Motivation, 50*, 307–338.

Gelman, A., & Hill, J. (2006). *Data analysis using regression and multilevel/hierarchical models*. New York, NY: Cambridge University Press.

Greene, J. D., Cushman, F. A., Stewart, L. E., Lowenberg, K., Nystrom, L. E., & Cohen, J. D. (2009). Pushing moral buttons: The interaction between personal force and intention in moral judgment. *Cognition, 111*, 364–371.

Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review, 108*, 814–834.

Haidt, J. (2003). The emotional dog does learn new tricks: A reply to Pizarro and Bloom (2003). *Psychological Review, 110*, 197–198.

Haidt, J. (2007). The new synthesis in moral psychology. *Science, 316*, 998–1002.

Haidt, J. (2008). Morality. *Perspectives on Psychological Science, 3*, 65–72.

Haidt, J., Bjorklund, F., & Murphy, S. (2000). *Moral dumbfounding: When intuition finds no reason.* Unpublished manuscript, University of Virginia.

Hall, L., Johansson, P., & Strandberg, T. (2012). Lifting the veil of morality: Choice blindness and attitude reversals on a self-transforming survey. *PloS One, 7*, e45457.

Hall, L., Johansson, P., Tärning, B., Sikström, S., & Deutgen, T. (2010). Magic at the marketplace: Choice blindness for the taste of jam and the smell of tea. *Cognition, 117*, 54–61.

Hall, L., Strandberg, T., Pärnamets, P., Lind, A., Tärning, B., & Johansson, P. (2013). How the polls can be both spot on and dead wrong: Using choice blindness to shift political attitudes and voter intentions. *PloS One, 8*, e60554.

Johansson, P., Hall, L., & Sikström, S. (2008). From change blindness to choice blindness. *Psychologia, 51*, 142–155.

Johansson, P., Hall, L., Sikström, S., & Olsson, A. (2005). Failure to detect mismatches between intention and outcome in a simple decision task. *Science, 310*, 116–119.

Kahane, G. (2012). On the wrong track: Process and content in moral psychology. *Mind & language, 27*, 519–545.

Knutson, K. M., Krueger, F., Koenigs, M., Hawley, A., Escobedo, J. R., Vasudeva, V., . . . Grafman, J. (2010). Behavioral norms for condensed moral vignettes. *Social Cognitive and Affective Neuroscience, 5*, 378–384.

Kohlberg, L. (1976). Moral stages and moralization: The cognitive developmental approach. In T. Lickona (Ed.), *Moral development and behavior: Theory, research and social issues* (pp. 31–53). New York, NY: Holt, Rinehart, & Winston.

Paxton, J. M., & Greene, J. D. (2010). Moral reasoning: Hints and allegations. *Topics in Cognitive Science, 2*, 511–527.

Pizarro, D. A., & Bloom, P. (2003). The intelligence of the moral intuitions: Comment on Haidt (2001). *Psychological Review, 110*, 193–196.

Pizarro, D. A., Uhlmann, E., & Bloom, P. (2003). Causal deviance and the attribution of moral responsibility. *Journal of Experimental Social Psychology, 39*, 653–660.

Royzman, E. B., Kim, K., & Leeman, R. F. (2015). The curious tale of Julie and Mark: Unraveling the moral dumbfounding effect. *Judgment and Decision Making, 10*, 296–313.

Royzman, E. B., Landy, J. F., & Leeman, R. F. (2015). Are thoughtful people more utilitarian? CRT as a unique predictor of moral minimalism in the dilemmatic context. *Cognitive Science, 39*, 325–352.

Sagana, A., Sauerland, M., & Merckelbach, H. (2014). Memory impairment is not sufficient for choice blindness to occur. *Frontiers in Psychology, 5*, 1–8.

Sauerland, M., Sagana, A., & Otgaar, H. (2013). Theoretical and legal issues related to choice blindness for voices. *Legal and Criminological Psychology, 18*, 371–381.

Trémolière, B., De Neys, W., & Bonnefon, J. F. (in press). Reasoning and moral judgment: A common experimental toolbox. In L. J. Ball & V. A. Thompson (Eds.), *International handbook of thinking and reasoning*. Hove, UK: Psychology Press.

Wheatley, T., & Haidt, J. (2005). Hypnotic disgust makes moral judgments more severe. *Psychological Science, 16*, 780–784.

Received October 14, 2015
Revision received March 4, 2016
Accepted March 4, 2016
Published online June 29, 2016

Štěpán Bahník
Department of Psychology
University of Würzburg
Röntgenring 10
97070 Würzburg
Germany
Tel. +49 931 318-0998
Fax +49 931 318-2812
E-mail bahniks@seznam.cz

